ABSTRACT

A system (520) generates labels for clusters of documents. The system (520) identifies topics associated with the documents in the clusters and determines whether the topics are associated with approximately half or more of the documents in the clusters. The system (520) then generates labels for the clusters using the topics that are associated with approximately half or more of the documents in the clusters.